



Plagiarism Checker X Originality Report

Similarity Found: 10%

Date: Saturday, June 28, 2025

Statistics: 252 words Plagiarized / 2403 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

Analisis Perbandingan Kinerja Algoritma Klasifikasi Data Menggunakan Metode K-NN, Naive Bayes, dan Decision Tree pada Dataset UCI Iris 1) Muhammad Dicky Azhary Octavianto Universitas Bina Darma, Jl. Jenderal Ahmad Yani No.3, 9/10 Ulu, Kecamatan Seberang Ulu I, Kota Palembang, Sumatera Selatan 30111 E-Mail: dickyzahari0210@gmail.com 2) Tata Subtari Universitas Bina Darma, Jl. Jenderal Ahmad Yani No.3, 9/10 Ulu, Kecamatan Seberang Ulu I, Kota Palembang, Sumatera Selatan 30111 E-Mail: tata.subtari@binadarma.ac.id ABSTRACT Data classification is one of the important techniques in data mining and machine learning, which is widely used to group data into certain classes.

This study aims to analyze and compare the performance of three classification algorithms, namely K-Nearest Neighbor (K-NN), Naive Bayes, and Decision Tree, in classifying Iris data from the UCI Machine Learning Repository. This dataset consists of 150 data with four feature attributes and three target classes. Testing was carried out using the cross-validation method with a k-fold approach of 10 folds.

The results of the performance evaluation were measured using the metrics of accuracy, precision, recall, and f1-score. Based on the test results, the K-NN algorithm showed the highest accuracy rate of 96.67%, followed by Decision Tree at 95.33%, and Naive Bayes at 94.00%. These findings indicate that choosing the right classification algorithm can affect the success rate in the data classification process.

Keyword : Data Mining, Classification, K-NN, Naive Bayes, Decision Tree, Iris Dataset

PENDAHULUAN Latar Belakang Dalam era digital saat ini, jumlah data yang dihasilkan dan disimpan terus meningkat secara signifikan. Pengolahan data tersebut menjadi informasi yang bernilai sangat penting bagi pengambilan keputusan di berbagai bidang. Salah satu teknik pengolahan data yang banyak digunakan adalah klasifikasi, yang merupakan bagian dari cabang ilmu data mining dan machine learning.

Klasifikasi digunakan untuk memetakan data ke dalam kategori tertentu berdasarkan karakteristik yang dimiliki oleh data tersebut. Berbagai algoritma klasifikasi telah dikembangkan, di antaranya adalah K-Nearest Neighbor (K-NN), Naive Bayes, dan Decision Tree. Masing-masing algoritma memiliki karakteristik dan performa yang berbeda tergantung pada jenis dan struktur data yang digunakan.

Oleh karena itu, penting untuk mengetahui bagaimana perbandingan kinerja masing-masing algoritma terhadap dataset tertentu. Dataset **Iris yang diperoleh dari UCI Machine Learning Repository** merupakan salah satu dataset standar yang sering digunakan dalam penelitian dan pembelajaran machine learning.

Dataset ini terdiri dari tiga jenis bunga Iris yang berbeda, yaitu Iris setosa, Iris versicolor, dan Iris virginica, dengan empat atribut fitur, yakni **panjang dan lebar sepal** serta **panjang dan lebar** petal. Dataset ini sangat ideal untuk melakukan studi perbandingan algoritma klasifikasi karena bersifat bersih, terstruktur, dan memiliki jumlah kelas yang seimbang.

Penelitian **ini bertujuan untuk membandingkan kinerja** tiga algoritma klasifikasi, yaitu K-NN, Naive Bayes, dan Decision Tree, dalam mengklasifikasikan data Iris. Evaluasi dilakukan dengan menggunakan metrik **akurasi, precision, recall, dan f1-score**, serta menerapkan teknik validasi silang agar hasil pengujian lebih objektif dan reliabel.

Rumusan Masalah Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah: Bagaimana kinerja masing-masing algoritma K-NN, Naive Bayes, dan Decision Tree dalam mengklasifikasikan data pada Iris Dataset? Algoritma mana yang memiliki tingkat akurasi dan efisiensi terbaik pada kasus klasifikasi data Iris? **Tujuan Penelitian** **Tujuan dari penelitian ini adalah** untuk: Menganalisis performa algoritma K-NN, Naive Bayes, dan Decision Tree dalam proses klasifikasi pada dataset Iris. Menentukan algoritma klasifikasi yang memiliki kinerja terbaik berdasarkan metrik evaluasi.

Manfaat Penelitian **Penelitian ini diharapkan dapat memberikan manfaat sebagai** berikut: Memberikan referensi kepada akademisi dan praktisi dalam memilih algoritma klasifikasi yang sesuai untuk jenis dataset tertentu. Menambah wawasan mengenai

perbandingan performa berbagai algoritma klasifikasi dalam konteks data yang sederhana dan terstruktur.

Mendorong penggunaan metode pengujian yang objektif dan sistematis dalam evaluasi algoritma machine learning. Ruang Lingkup Penelitian Penelitian ini dibatasi hanya pada tiga algoritma klasifikasi, yaitu K-NN, Naive Bayes, dan Decision Tree. Dataset yang digunakan adalah Iris Dataset dari UCI yang terdiri dari 150 data.

Evaluasi dilakukan berdasarkan metrik **akurasi, precision, recall, dan f1-score** menggunakan metode 10-fold cross-validation. Penelitian tidak melibatkan data primer maupun interaksi langsung dengan responden. TINJAUAN PUSTAKA Data Mining dan Klasifikasi Data mining merupakan proses penemuan pola atau pengetahuan yang tersembunyi dari data dalam jumlah besar.

Salah satu tugas utama dalam data mining adalah klasifikasi, yaitu proses untuk memetakan data ke dalam kelas-kelas yang telah ditentukan sebelumnya berdasarkan atribut-atribut yang dimiliki oleh data tersebut (Han, Kamber & Pei, 2012). Klasifikasi termasuk dalam pembelajaran terawasi (supervised learning), karena memerlukan data latih yang sudah diberi label kelas untuk membangun model.

Model ini kemudian **digunakan untuk memprediksi kelas dari data baru yang** belum diketahui kelasnya. Dataset Iris Dataset Iris adalah dataset yang dikumpulkan oleh Ronald A. Fisher pada tahun 1936 dan kini tersedia secara publik di **UCI Machine Learning Repository**. Dataset ini terdiri dari tiga jenis bunga Iris: **Setosa, Versicolor, dan Virginica** dengan empat atribut fitur, yaitu: Sepal Length (cm) Sepal Width (cm) Petal Length (cm) Petal Width (cm) **Dataset ini sering digunakan** sebagai benchmark dalam pengujian dan pembelajaran algoritma klasifikasi.

Algoritma K-Nearest Neighbor (K-NN) K-NN adalah algoritma klasifikasi berbasis instance-based learning yang bekerja dengan prinsip mencari sejumlah tetangga terdekat (nilai k) dari data baru untuk menentukan kelasnya. Jarak antar data umumnya diukur menggunakan Euclidean distance. K-NN tidak memerlukan proses pelatihan model dan bersifat non-parametrik, namun performanya sangat tergantung pada nilai k dan distribusi data (Cover & Hart, 1967).

Algoritma Naive Bayes **Naive Bayes adalah algoritma klasifikasi** probabilistik **yang didasarkan pada Teorema** Bayes, dengan asumsi bahwa setiap atribut adalah independen terhadap atribut lainnya. Meskipun asumsi ini seringkali tidak sepenuhnya benar dalam data dunia nyata, Naive Bayes tetap memberikan performa yang baik pada banyak kasus klasifikasi.

Keunggulan utamanya terletak pada kecepatan dan efisiensinya dalam menangani data berskala besar (Rish, 2001). Algoritma Decision Tree **Decision Tree adalah metode** klasifikasi berbasis pohon keputusan, di mana data dibagi secara rekursif berdasarkan atribut tertentu yang memberikan informasi terbaik (misalnya menggunakan gain, entropy, atau Gini index). Algoritma yang umum digunakan dalam Decision Tree antara lain ID3, C4.5, dan CART.

Model Decision Tree **mudah dipahami dan diinterpretasikan** (Quinlan, 1996). Penelitian Terkait Beberapa penelitian sebelumnya telah membandingkan algoritma-algoritma klasifikasi menggunakan dataset Iris. Misalnya, penelitian oleh Prasetyo (2020) menunjukkan bahwa algoritma K-NN memiliki akurasi lebih tinggi dibanding Naive Bayes dan Decision Tree pada Iris Dataset.

Penelitian lain oleh Lestari & Wijaya (2021) menemukan bahwa validasi silang (cross-validation) mampu meningkatkan akurasi model klasifikasi dengan hasil terbaik pada K-NN dengan $k=5$. Studi-studi ini menunjukkan bahwa meskipun dataset yang digunakan sama, hasil performa algoritma dapat bervariasi tergantung pada metode evaluasi, preprocessing, dan parameter yang digunakan.

METODOLOGI PENELITIAN Metode Penelitian Penelitian ini menggunakan metode eksperimen komputasi, yaitu dengan menerapkan tiga algoritma klasifikasi pada dataset publik (Iris Dataset) untuk dianalisis performanya. Pengujian dilakukan secara objektif menggunakan metrik evaluasi standar, tanpa melibatkan pengumpulan data primer atau interaksi dengan responden.

Alur Penelitian Adapun tahapan proses penelitian ini dapat digambarkan dalam diagram berikut: / Gambar 1. Alur Penelitian Diagram alur penelitian terdiri dari beberapa tahapan utama, yang dijelaskan sebagai berikut: Mulai (Start) Proses penelitian diawali dengan perencanaan dan penentuan tujuan penelitian, yaitu membandingkan performa tiga algoritma klasifikasi. Pengambilan Dataset Dataset Iris diunduh **dari UCI Machine Learning Repository**.

Dataset ini berisi data spesies bunga Iris dengan 150 entri dan empat fitur numerik. Preprocessing Data Pada tahap ini, dilakukan normalisasi data (menggunakan Min-Max Normalization), pemeriksaan data duplikat atau kosong, serta encoding kelas target jika diperlukan. Pembagian Data (10-Fold Cross Validation) Dataset dibagi ke dalam 10 bagian (fold).

Secara bergantian, 9 bagian digunakan **sebagai data latih dan 1** bagian sebagai data uji.

Proses ini diulang sebanyak 10 kali untuk memperoleh evaluasi yang lebih stabil. Implementasi Algoritma Tiga algoritma diterapkan pada data: K-NN: Menggunakan parameter $k = 3, 5, \text{ dan } 7$ untuk menentukan tetangga terdekat.

Naive Bayes: Menggunakan distribusi Gaussian untuk menghitung probabilitas. Decision Tree: Menggunakan metode CART dengan Gini Index sebagai pengukur impurity. Evaluasi Kinerja Hasil klasifikasi dari tiap algoritma dievaluasi menggunakan metrik: Akurasi Precision Recall F1-Score Evaluasi berdasarkan confusion matrix dari tiap fold, lalu dirata-rata. Analisis Hasil Dibandingkan performa antar algoritma untuk menarik kesimpulan algoritma mana yang paling optimal dalam kasus ini.

Selesai (End) Penelitian ditutup dengan penyusunan laporan, kesimpulan, dan rekomendasi untuk penelitian selanjutnya.. Akuisisi dan Persiapan Data Dataset Iris diperoleh dari situs <https://archive.ics.uci.edu/ml/datasets/iris> . Dataset ini memiliki 150 baris data dan 5 kolom (4 atribut fitur + 1 label kelas).

Data dibagi menjadi dua bagian: Data latih dan uji menggunakan metode 10-Fold Cross Validation agar evaluasi lebih akurat dan menghindari overfitting. Data juga dilakukan normalisasi menggunakan Min-Max Normalization agar setiap fitur memiliki skala yang seragam antara 0 hingga 1, khusus untuk algoritma K-NN. Implementasi Algoritma Tiga algoritma yang diimplementasikan adalah: K-Nearest Neighbor (K-NN) Parameter: nilai k ditentukan optimal secara eksperimen (misal: 3, 5, 7). Metode pengukuran jarak: Euclidean Distance. Naive Bayes Menggunakan distribusi Gaussian untuk atribut numerik. Tidak memerlukan parameter tuning.

Decision Tree Algoritma: CART (Classification and Regression Tree). Kriteria pemisahan: Gini Index. Tabel 1. Parameter Algoritma Algoritma _Parameter Utama _Catatan _K-NN _ $k = 3, 5, 7$ _Pengaruh besar terhadap akurasi _Naive Bayes _ _Asumsi distribusi Gaussian _ _Decision Tree _Max Depth (opsional) _Menggunakan Gini Index _ _Evaluasi Kinerja Evaluasi dilakukan menggunakan metrik: Akurasi = (Jumlah Prediksi Benar) / (Total Data) Precision = $TP / (TP + FP)$ Recall = $TP / (TP + FN)$ F1-Score = $2 \times (Precision \times Recall) / (Precision + Recall)$ Di mana: TP: True Positive FP: False Positive FN: False Negative Evaluasi dilakukan dengan 10-Fold Cross Validation agar setiap algoritma diuji pada data yang berbeda secara adil.

Tools dan Perangkat Penelitian ini dilakukan dengan menggunakan: Bahasa pemrograman Python Library: scikit-learn, pandas, numpy, matplotlib, seaborn Komputasi dijalankan pada PC/Laptop standar (Windows 10, RAM 8GB) HASIL DAN PEMBAHASAN Hasil Evaluasi Setelah dilakukan pengujian terhadap tiga algoritma klasifikasi menggunakan metode 10-Fold Cross Validation, diperoleh hasil sebagai

berikut: Tabel 2. Hasil Evaluasi Kinerja Algoritma.

Algoritma _Akurasi (%) _Precision _Recall _F1-Score _K-NN (k=3) _96.67 _0,067361
_0,067361 _0,066667 _Naive Bayes _94.00.00 _0,065278 _0,065278 _0,065278 _
_Decision Tree _95.33.00 _0,065972 _0,065972 _0,065972 _ _ / Gambar 2. Grafik
Perbandingan Akurasi Algoritma Analisis Hasil Berdasarkan tabel dan grafik di atas,
dapat disimpulkan bahwa: K-NN memperoleh **nilai akurasi tertinggi yaitu 96.67%**.

Hal ini disebabkan karena metode K-NN sangat cocok diterapkan pada dataset kecil dan seimbang seperti Iris, serta efektif dalam menangani data numerik dengan normalisasi yang tepat. Hasil dari eksperimen menunjukkan bahwa algoritma K-NN memiliki performa tertinggi dalam hal akurasi klasifikasi pada dataset Iris. Hal ini dapat dijelaskan karena K-NN mempertimbangkan kedekatan data dalam ruang multidimensi, yang sangat cocok untuk dataset kecil dengan distribusi yang jelas seperti Iris.

Sifat non-parametrik K-NN memungkinkan algoritma ini menyesuaikan terhadap bentuk distribusi data yang tidak linier Decision Tree menempati posisi kedua dengan akurasi 95.33%, menunjukkan kinerja yang stabil dan mudah diinterpretasikan. **Keunggulan utama dari Decision Tree adalah** kemampuannya untuk menghasilkan model yang mudah dipahami secara visual serta fleksibel dalam menangani data numerik dan kategorikal. Namun, Decision Tree cenderung mengalami overfitting, terutama jika tidak dilakukan pruning atau jika data memiliki noise.

Naive Bayes memperoleh akurasi 94.00%, sedikit lebih rendah dibanding dua algoritma lainnya. Hal ini kemungkinan disebabkan oleh **asumsi independensi antar fitur** yang tidak sepenuhnya terpenuhi dalam Iris Dataset. Meskipun sederhana dan efisien secara komputasi, **asumsi independensi antar fitur** yang dimiliki Naive Bayes sering kali tidak realistis, terutama pada dataset dengan korelasi antar variabel.

Namun, algoritma ini tetap relevan dan kuat dalam kasus tertentu yang sesuai dengan asumsi dasarnya. Pembahasan Komparatif Setiap algoritma memiliki keunggulan dan kelemahannya masing-masing: K-NN unggul dalam akurasi, namun memerlukan komputasi yang tinggi saat prediksi jika dataset besar. Naive Bayes sangat cepat dan ringan secara komputasi, namun keakuratannya sensitif terhadap asumsi independensi fitur.

Decision Tree sangat interpretatif dan fleksibel, namun rentan terhadap overfitting jika tidak dibatasi (misal dengan pruning). Kombinasi antara akurasi dan efisiensi menjadikan K-NN sebagai algoritma terbaik dalam konteks penelitian ini. Namun dalam skala besar atau data tidak seimbang, hasil bisa berbeda.

Perbandingan ini menunjukkan bahwa tidak ada satu algoritma yang secara mutlak lebih unggul di semua kasus. Pemilihan algoritma terbaik sangat bergantung pada karakteristik data yang digunakan, ukuran dataset, serta kebutuhan akan interpretabilitas model. Untuk dataset kecil dan bersih seperti Iris, K-NN memberikan hasil terbaik; namun untuk skenario lain, algoritma lain mungkin lebih sesuai.

Selain akurasi, metrik precision, recall, dan f1-score juga menunjukkan tren serupa, memperkuat validitas temuan bahwa K-NN memberikan hasil yang konsisten lebih baik. Penelitian ini memberikan gambaran awal yang dapat menjadi dasar pemilihan algoritma dalam proyek klasifikasi data serupa. KESIMPULAN DAN SARAN Kesimpulan Berdasarkan hasil pengujian dan analisis pada penelitian ini, dapat disimpulkan bahwa: Algoritma K-Nearest Neighbor (K-NN) memberikan performa terbaik dalam klasifikasi data Iris dengan akurasi tertinggi sebesar 96,67%, diikuti oleh Decision Tree dengan akurasi 95,33%, dan Naive Bayes sebesar 94,00%.

Evaluasi menggunakan metrik precision, recall, dan f1-score juga menunjukkan hasil konsisten, di mana K-NN memiliki nilai tertinggi pada setiap metrik. Karakteristik dataset seperti ukuran, keseimbangan antar kelas, dan bentuk fitur sangat memengaruhi kinerja dari masing-masing algoritma. Pemilihan algoritma klasifikasi yang tepat perlu disesuaikan dengan kondisi data, tujuan penggunaan, serta efisiensi komputasi yang diharapkan.

Saran Penelitian ini dapat dikembangkan lebih lanjut dengan saran berikut: Menggunakan dataset lain yang lebih kompleks, tidak seimbang, atau berjumlah besar untuk melihat stabilitas performa algoritma. Menerapkan teknik hybrid atau ensemble learning seperti Random Forest atau Voting Classifier untuk meningkatkan akurasi. Menambahkan tahapan feature selection atau dimensionality reduction seperti PCA untuk meningkatkan efisiensi klasifikasi. DAFTAR PUSTAKA Han, J., Kamber, M.,

& Pei, J. (2012). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. Rish, I. (2001). An empirical study of the naive Bayes classifier. IJCAI 2001 Workshop on Empirical Methods in AI. Quinlan, J. R. (1996). Improved use of continuous attributes in C4.5.

Journal of Artificial Intelligence Research, 4, 77–90. UCI Machine Learning Repository. (n.d.). Iris Data Set. Retrieved from <https://archive.ics.uci.edu/ml/datasets/iris> Prasetyo, E. (2020). Data Mining: Konsep dan Aplikasi Menggunakan MATLAB. Andi. Lestari, R., & Wijaya, H. (2021). Perbandingan algoritma klasifikasi pada dataset Iris. Jurnal Teknologi

dan Informatika, 15(2), 125–134. Witten, I. H., Frank, E., & Hall, M. A. (2016).

Data Mining: Practical Machine Learning Tools and Techniques (4th ed.). Morgan Kaufmann. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.

Tan, P. N., Steinbach, M., & Kumar, V. (2019). **Introduction to Data Mining** (2nd ed.).

Pearson. Rokach, L., & Maimon, O. (2014). Data Mining with Decision Trees: Theory and Applications. World Scientific. Zhang, H. (2004). The optimality of naive Bayes. FLAIRS Conference, 1, 563–567.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). **An Introduction to Statistical Learning**. Springer. Raschka, S. (2015). Python Machine Learning. Packt Publishing.

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

INTERNET SOURCES:

<1% -

<https://journal.darmajaya.ac.id/index.php/JoDMAApps/article/download/911/265/3519>

<1% - <https://bpm.binadarma.ac.id/dokumen/detail/6/standar-spmi>

<1% -

<https://medium.com/@aliwasim.12aaa64/classification-in-data-mining-a-complete-guide-df9239f41ba8>

<1% -

<https://pdfs.semanticscholar.org/cb00/ed07e9189357ea7f8612738b251f8db61d37.pdf>

<1% - <https://machinelearning.org.in/accuracy-precision-recall-and-f1-score/>

<1% - <https://www.nature.com/articles/s41598-024-76909-6>

<1% - <https://dataaspirant.com/classification-algorithms/>

<1% -

<https://id.scribd.com/presentation/720180324/4-Algoritma-Klasifikasi-K-Nearest-Neighbor>

<1% - <http://ojsamik.amikmitragama.ac.id/index.php/js/article/download/209/197>

1% - <https://conference.upnvj.ac.id/index.php/senamika/article/view/3125>

<1% -

https://www.academia.edu/108071654/Analisis_Perbandingan_Kinerja_Algoritma_Klasifikasi_dengan_Menggunakan_Metode_K_Fold_Cross_Validation

<1% - <https://jurnal.darmajaya.ac.id/index.php/PSND/article/download/2948/1267>

<1% - <https://id.scribd.com/doc/195393571/Kegunaan-Penelitian>

<1% -

<https://medium.com/@basriumar/part-13-evaluasi-model-metode-dan-metrik-penting-dalam-machine-learning-a367db87696e>

<1% -

<https://skripsiuk.com/teknik-pengumpulan-data-fondasi-utama-penelitian-yang-akurat-dan-terpercaya/>

<1% -

<https://www.kompasiana.com/hadijahnisaifayatin/632f45f24addee4cd153ac06/jenis-atribut-karakteristik-tipe-dan-teknik-data-mining>

<1% - <https://juti.if.its.ac.id/index.php/juti/article/download/384/273/0>

<1% - <https://dqlab.id/pelajari-3-data-set-machine-learning-bersama-dqlab>

<1% - <https://id.scribd.com/document/494304082/Data-Iris>

<1% -

<https://sis.binus.ac.id/2025/01/15/apa-itu-naive-bayes-algoritma-klasifikasi-sederhana-yang-kuat-dalam-machine-learning/>

<1% -

<https://journal3.uin-alauddin.ac.id/index.php/teknosains/article/download/38383/19022/>

<1% - <https://ilmukomputer.org/wp-content/uploads/2019/06/Rina-Algoritma-C45.pdf>

<1% - <https://repository.unmuhjember.ac.id/7131/1/JURNAL.pdf>

<1% - https://eprints.unisnu.ac.id/id/eprint/1063/5/171240000725_BAB%20IV.pdf

<1% -

<https://dibimbing.id/blog/detail/memahami-teknik-validasi-silang-k-fold-dalam-machine-learning>

<1% -

<https://pdfs.semanticscholar.org/389c/5266c625b11038a94b7a98222da1661daed0.pdf>

<1% - <https://www.kaggle.com/datasets/dskagglemt/iris-uci-dataset>

<1% -

<https://rumahcoding.id/blog/normalisasi-dan-standarisasi-data-dengan-scikit-learn/>

<1% -

<https://repository.uin-suska.ac.id/75481/1/UPLOAD%20JURNAL%20REPOSITORY.pdf>

<1% - <https://pub.nuris.ac.id/sainsdata/article/download/31/30>

<1% -

https://www.academia.edu/57586902/Analisis_Perbandingan_Algoritma_Decision_Tree_k_NN_dan_Naive_Bayes_untuk_Prediksi_Kesuksesan_Start_up

<1% -

<https://ruangjurnal.com/mengenal-model-decision-tree-konsep-implementasi-dan-aplikasi-dalam-dunia-data-science/>

<1% -

<https://www.karirinfo.com/2024/11/perbandingan-algoritma-svm-knn-decision-tree-naive-bayes.html>

<1% - <https://digilib.uinsgd.ac.id/101519/>

<1% -

<https://www.gudanginformatika.com/2025/01/memahami-algoritma-pemrograman-dengan.html>

<1% - <https://id.scribd.com/document/881293995/Laporan-Praktikum-5-Naive-Bayes>

<1% - https://elibrary.unpas.ac.id/index.php?p=show_detail&id=14696

<1% - <https://www.scirp.org/reference/referencespapers?referenceid=1263126>

1% - <https://sci-hub.se/10.1007/978-1-4614-7138-7>